An Initial-Fit Comparison of Two Generic Hearing Aid Prescriptive Methods (NAL-NL2 and CAM2) to Individuals Having Mild to Moderately Severe High-Frequency Hearing Loss

Earl E. Johnson*†

Abstract

Background: Johnson and Dillon (2011) provided a model-based comparison of current generic hearing aid prescriptive methods for adults with hearing loss based on the attributes of speech intelligibility, loudness, and bandwidth.

Purpose: This study compared the National Acoustic Laboratories—Non-linear 2 (NAL-NL2) and Cambridge Method for Loudness Equalization 2—High-Frequency (CAM2) prescriptive methods using adult participants with less high-frequency hearing loss than Johnson and Dillon (2011). Of study interest was quantification of prescribed audibility, speech intelligibility, and loudness. The preferences of participants for either NAL-NL2 or CAM2 and preferred deviations from prescribed settings are also reported.

Research Design: Using a single-blind, counter-balanced, randomized design, preference judgments for the prescriptive methods with regard to sound quality of speech and music stimuli were obtained. Preferred gain adjustments from the prescription within the 4–10 kHz frequency range were also obtained from each participant. Speech intelligibility and loudness model calculations were completed on the prescribed and adjusted amplification.

Study Sample: Fourteen male Veterans, whose average age was 65 yr and whose hearing sensitivity averaged normal to borderline normal through 1000 Hz sloping to a moderately severe sensorineural loss, served as participants.

Data Collection and Analysis: Following a brief listening time (~10 min), typical of an initial fitting visit, the participants made paired comparison of sound quality between the NAL-NL2 and CAM2 prescriptive settings. Participants were also asked to modify each prescription in the range of 4–10 kHz using an overall gain control and make subsequent comparisons of sound quality preference between prescriptive and adjusted settings. Participant preferences were examined with respect to quantitative analysis of loudness modeling, speech intelligibility modeling, and measured high-frequency bandwidth audibility.

Results: Consistent with the lack of difference in predicted speech intelligibility between the two prescriptions, sound quality preferences on the basis of clarity were split across participants while some participants did not have a discernable preference. Considering sound quality judgments of pleasantness, the majority of participants preferred the sound quality of the NAL-NL2 (8 of 14) prescription instead of the CAM2 prescription (2 of 14). Four of the 14 participants showed no preference on the basis of pleasantness for either prescription. Individual subject preferences were supported by loudness modeling that indicated NAL-NL2 was the softer of the two prescriptions and CAM2 was the louder. CAM2 did provide more audibility to the higher frequencies (5–8 kHz) than NAL-NL2. Participants turned the 4–10 kHz gain

---

*James H. Quillen VA Medical Center, Mountain Home, TN; †Department of Audiology and Speech-Language Pathology, East Tennessee State University, Johnson City, TN

Earl E. Johnson, Mountain Home VA Medical Center, PO Box 4000, Audiology-126, Mountain Home, TN 37684; Phone: 423-926-1171 x7149; Fax: 423-979-3403

Disclaimer: The opinions expressed in this article are those of the authors and do not necessarily represent the official position of the U.S. Department of Veterans Affairs or the United States government.

recommendation of CAM2 lower, on average, by a significant amount of 4 dB when making adjustments while no significant adjustment was made to the initial NAL-NL2 recommendation.

Conclusions: NAL-NL2 prescribed gains were more often preferred at the initial fitting by the majority of participating veterans. For those patients with preference for a louder fitting than NAL-NL2, CAM2 is a good alternative. When the participant adjustment from the prescription between 4 and 10 kHz exceeded 4 dB from either NAL-NL2 (2 of 14) or CAM2 (11 of 14), the participants demonstrated a later preference for that adjustment 69% of the time. These findings are viewed as limited evidence that some individuals may have a preference for high-frequency gain that differs from the starting prescription.

Key Words: Adjustment, audibility, CAM2, CAMEQ2-HF, gain, hearing aid, loudness, NAL-NL2, preference, prescription, sound quality, speech intelligibility, training

Abbreviations: CST = Connected Speech Test; FFT = fast Fourier transformation; IG = insertion gain; ILTASS = International Long-Term Average Speech Spectrum; NAL-NL1 = National Acoustic Laboratories—Non-linear 1; NAL-NL2 = National Acoustic Laboratories—Non-linear 2; REAR = real-ear aided response; RMS = root mean square; SII = Speech Intelligibility Index

Many hearing aids adjust their response characteristics over time, referred to generally as training, on the basis of adjustments made by users. This capability, however, still relies on the initial prescribing of hearing aid amplification characteristics for individual listeners. Research has indicated when listeners are allowed to train the frequency- and input-dependent gain of the hearing aids, the end result is affected by the starting response (baseline) (Dreschler et al, 2008; Keidser et al, 2008; Mueller et al, 2008). That is, the final trained setting of the hearing aid depends on the starting point of training. Mueller et al (2008) used an analysis of effect size, partial eta squared η^2, to determine that 92% of the variability in preferred gain deviations was attributable to the starting point. Keidser et al (2008) further indicated that some trained settings are not overwhelmingly "preferred" settings compared to untrained settings, and Dreschler et al (2008) indicated the end trained response is affected by the training interface device itself.

Practitioners have generally been cautious about basing amplification parameters on patient preference; that is, preferred amplification parameters for highly regarded sound quality are not necessarily the ones that lead to optimal performance on the dimension of speech intelligibility (Logan et al, 1984; Harris and Goldstein, 1985). Albeit including patient preferences is an integral part of the clinic service delivery model encouraging an interactive and patient-driven, practitioner-facilitated exchange, a patient is also intuitively much more likely to wear a hearing aid with the amplification parameters he or she prefers (e.g., Humes and Hackett, 1986; Smeds et al, 2006a, 2006b; Johnson and Dillon, 2011).

Today, the core of modern hearing aid fittings is still the prescriptive method that, by no small coincidence, has been based on the pure-tone hearing loss thresholds, an omnipresent form of hearing assessment. Pure tone audiometry thresholds also remain the primary mode of hearing assessment for modeling loudness and speech intelligibility results across a large diverse group of hearing-impaired listeners (American National Standards Institute [ANSI], 1997; Moore and Glasberg, 2004; Dillon et al, 2008). The highly regarded ANSI S3.5 (ANSI, 1997) Speech Intelligibility Index (SII) model and the Moore and Glasberg (2004) loudness model thus require pure tone thresholds as the sole input variable to represent hearing ability rather than, or in addition to, other domains of hearing assessment such as frequency or temporal resolution.

Two of three principal components pertinent to modeling hearing aid outcomes, that is, speech recognition performance and subject benefit/satisfaction, have been best accounted for by speech intelligibility and loudness estimates, respectively (Humes, 2003). Prescriptive approaches for hearing aids have thus been careful to recommend amplification gain with consideration to predicted speech intelligibility and/or loudness (Rankovic, 1991; Cornelisse et al, 1995; Cox, 1995; Byrne et al, 2001; Moore et al, 2010; Keidser et al, 2011). Johnson and Dillon (2011) reported on four prominent generic hearing aid prescriptions indicating that all were comparable with regard to predicted speech intelligibility but did differ markedly in terms of predicted loudness. The Johnson and Dillon (2011) study focused on generic prescriptive procedures because they are not specific to, and may be used with, any brand or product model of hearing aid.

Two validated generic prescriptive approaches for hearing aids are the National Acoustic Laboratories—Non-linear 2 method (NAL-NL2) (Keidser et al, 2011) and the Cambridge Method for Loudness Equalization 2—High-Frequency (CAM2) (Moore et al, 2010; Moore and Füllgrabe, 2010; Moore et al, 2011). Both prescriptive methods are used for adult hearing aid fittings throughout the world with varying popularity in different countries and regions. In practice, the hearing aid fitting practitioner typically chooses the single prescription that he or she will use almost exclusively with their patients. A prescriptive approach may be selected via the programming software (a more passive approach) and/or

through verification of real-ear targets (a more active approach), which is the best-practice recommendation (Valente et al, 2006; Kochkin, 2011).

Besides CAM2 and NAL-NL2 other prescriptions such as the Desired Sensation Level (DSL) m[i/o] 5.0 (Scollie et al, 2005) and brand-specific prescriptions are available to practitioners and patients; however, this study only evaluated NAL-NL2 and CAM2 because they are based on similar speech intelligibility and loudness modeling themes. The NAL-NL2 and CAM2 prescriptions are both loudness equalization approaches as opposed to loudness normalization approaches, which had some popularity one to two decades ago (for review, see Ricketts and Van Vliet, 1996).

The loudness equalization approach means that the perception of loudness is equivalent or near equivalent across pitches, generally from approximately 500 to 4000 Hz, at the amplified levels. This is in contrast with individuals having normal hearing where the loudness perception is dominated by the lower frequencies around 500 Hz (e.g., Fletcher and Munson, 1933; Moore and Glasberg, 2004; Johnson and Dillon, 2011). In contrast, the philosophy of loudness normalization used by past prescriptive methods, such as the Independent Hearing Aid Fitting Forum (IHAFF) (Cox, 1995; Valente and Van Vliet, 1997) and Figure 6 (Killion, 1994; Killion 1995), sought to create amplified loudness across soft to loud input levels that approximates the perception of the typical normal hearer in lieu of equalizing loudness across a frequency range. The DSL family of prescriptions is a current-day example of a loudness normalization approach (Cornelisse et al, 1995; Scollie et al, 2005), although it has been reported that slight gain modifications can offer a degree of loudness equalization (Seewald et al, 1996).

Similar to National Acoustic Laboratories—Nonlinear 1 (NAL-NL1), the NAL-NL2 prescriptive method maintained the rationale of maximizing speech intelligibility rationale within a constrained loudness budget allocated equally across frequencies. NAL-NL2 was based on a new SII model incorporating an internally revised desensitization factor and the loudness model of Moore and Glasberg (2004) for a possible prescribed bandwidth up to 10 kHz (Dillon et al, 2008; Keidser et al, 2011). Generally speaking when compared to NAL-NL1, NAL-NL2 prescribes less gain for adults, more gain for children, and larger compression ratios for both adults and children with mild and moderate hearing loss, and restricts the ratio of compression (e.g., less than 2:1) for those with severe/profound hearing loss. NAL-NL2 also introduced patient-specific modifications for gain adaptation including hearing aid experience, gender, and tonality of the patient-spoken language.

CAM2 aims to place as much of the speech spectrum as possible above absolute threshold for a given overall loudness (Moore et al, 2010). A defining feature of CAM2 is its prescription of gain at 10 kHz to place the root-mean-square (RMS) level of speech at this frequency equal to the hearing threshold level no matter the severity of hearing loss. As an example, to partially restore audibility for a hearing loss of 65 dB HL at 10 kHz would require an effective insertion gain of about 36 dB (Moore et al, 2008). Sample validation studies with CAM2 include Füllgrabe et al (2010), Moore and Füllgrabe (2010), and Moore et al (2011); recently, a gain adjustment for new versus experienced hearing aid users has been included that was not available at the time the current study was completed (Moore, pers. comm.).

Because practitioners tend to use the same prescriptive approach for the majority of their patients, the study completed is regarded as useful for the creation of initial evidence comparing the two prominent prescriptions. It is expected that practitioners are interested in the gain, speech intelligibility, and loudness differences between the two prescriptions as well as the initial preference that patients may exhibit or not exhibit for one or the other. Also, because training the amplification parameters of a hearing aid is now commonplace for input-dependent gain and frequency response, studying initial preference at prescriptive starting points is important because there is less of a necessity for patients to acclimatize to fixed amplification settings.

METHODS

Participants

Fourteen participants with mild to moderately severe sensorineural hearing losses enrolled at the Mountain Home Veterans Affairs Medical Center (Fig. 1). Eight participants were experienced hearing aid users of 6 mo or longer, and six participants were new users; all participants were male. The average participant age was 65 yr with a standard deviation of 5 yr.

Stimuli

Three test stimuli materials were used for paired comparisons of sound quality between the generic prescriptive methods: female speech in quiet, female speech in noise, and classical music. For the paired comparison in quiet, two sentences from the Connected Speech Test (CST) (Cox et al, 1987) of ~5 sec total duration were presented with the amplification of one prescription followed by ~1 sec of silence and then followed by two more CST sentences (~5 sec with the amplification of the other prescription. For the paired comparison of speech in noise, sentences from the CST were presented in the presence of recorded cafeteria
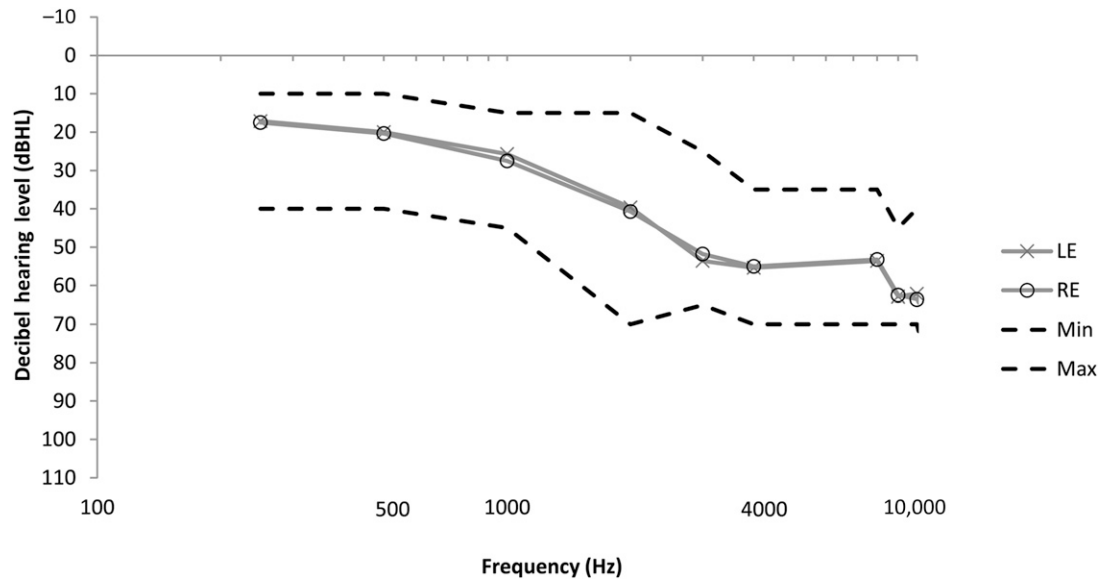
Figure 1. Average audiogram of the 14 participants and the minimum and maximum thresholds of any particular participant.

noise (Auditec, Inc., St. Louis, MO) at a 10 dB signal-to-noise ratio. The CST speech material was shaped to the International Long-Term Average Speech Spectrum (ILTASS) of Byrne et al (1994) with a Blackmann windowing function and fast Fourier transformation (FFT) size of 2048 using Adobe Audition 2.0. The accuracy of shaping was verified using a Larson Davis 824 sound level meter in one-third octave bands to within an average of 0.58 dB between 0.160 and 10 kHz (range of +1.5 dB at 6.3 kHz and −1.2 dB at 8 kHz) against the Byrne et al (1994) ILTASS publication. A 5 sec excerpt from Bach's *Well-Tempered Clavier Book Two,* Prelude and Fugue in C Minor, served as the music stimuli for comparison between prescriptions.

Stimuli material was streamed from a desktop computer employing Adobe Audition version 2.0, which routed externally through a MOTU (Miracle of the Unicorn) 24 input/output sound card, to a Tucker Davis Systems Technology System 3 running the Windows-based software application *FreqShaping* employing linear gain shaping by frequency. The stimuli material was output through Etymotic Research headphones (ER-2) and ER1-14A ear foam tips. Because the focus was on average level speech, nonlinear (compression) amplification was not a necessity because nonlinear and linear gain amplifiers are designed to provide the same amount of gain to average level speech input (e.g., Byrne, 1996; Byrne, et al, 2001; Dillon, 2001). ER-2 headphones were chosen because of their flat response, particularly through the higher frequencies up to 10 kHz. The *FreqShaping* program, developed by Peter Yang in 2004, has been used in experimental studies with the hearing impaired such as Hornsby et al (2011). This program was chosen because of the expected ability to better fit the high-frequency gain

recommendations of CAM2 provided the report of difficulty fitting CAM2 gain recommendations above 7.5 kHz using commercially available hearing aids (e.g., Füllgrabe et al, 2010).

Also, to fit the recommended high-frequency targets of CAM2, a determination was made to fit in a closed canal condition, which would lessen the likelihood of feedback occurrence. This is despite the fact that in perhaps most clinical environments, hearing losses of the magnitude included would be fit in an "open" ear canal configuration where achieving a prescriptive target may be regarded with less importance. That is, to compare the NAL-NL2 and CAM2 prescriptive methods, the ability to fit the prescriptive targets of each method was of paramount consideration.

Insertion Gain Fittings

For verification of the recommended targets of each prescriptive method, an insertion gain approach using a Knowles Electronic Manikin for Acoustic Research (KEMAR) was completed for several reasons. First, there are currently no commercially available verification units available for real-ear measures that include the recently introduced CAM2 and NAL-NL2 fitting methods. Secondly, attempting to verify beyond 6 kHz with the conventional probe-microphone equipment and approach is challenging. It is challenging because no commercially available real-ear probe microphone equipment extends beyond 8 kHz, but both prescriptive releases make target recommendations through 10 kHz. A second challenge relates to proximity of the probe tube to the eardrum when recording accurate sound pressure levels in the 6–8 kHz range. For example, to record at 6 and 8 kHz using a probe tube within a

141

−2 dB accuracy difference of dB measured at the eardrum, the probe tube must be within ∼7.5 and ∼5 mm of the eardrum, respectively (Dirks and Kincaid, 1987). Hence, an approach to verification that did not rely on probe microphone measures was devised.

In the sound field at a level of 65 dBA SPL from a loudspeaker (a Genelec model 1031A) 1 m away, the ILTASS shaped calibration noise of the CST material was presented to the open ear of KEMAR fit with IEC 711 simulated ear canal couplers. Terminating at the end of the canal opposite the open ear was the RA0045 microphone diaphragm, also referred to as the eardrum substitute. The microphone was routed to a GRAS 12AA two-channel power module onto a Larson Davis 824 sound level meter via the ADP005 B-N-C to GRAS type 26 AC 1/4 inch preamplifier attachment. This enabled the measurement of an unaided KEMAR response. An aided KEMAR response was measured in similar fashion upon the placement of ER-2 headphones and ER1-14A ear foam tips in the 711 coupler (KEMAR ear canal). The measured difference between the unaided and aided response was regarded as the obtained insertion gain. A separate insertion gain fit to target was obtained for each of the 14 subjects and both prescriptions.

Excellent NAL-NL2 and CAM2 fittings were obtained to insertion gain targets though 8 kHz (Fig. 2) similar to that achieved by Füllgrabe et al (2010) for CAM2 targets. Because of the steep climb in insertion gain required by CAM2 at 10 kHz, there was not adequate discrete gain control by frequency to fit targets from 5 to 8 kHz as well as 10 kHz. Nevertheless, these fittings through 8 kHz are as good as can be obtained on all currently available real-ear verification equipment,

which only verify through 8 kHz as well. Figure 3 displays the average difference between CAM2 and NAL-NL2 in their initial prescribed insertion gain by frequency response for the participants as well as the obtained insertion gain frequency response.

The participants, with normal typanograms, including peak pressure, compliance, and ear canal volume, were substituted for KEMAR following the insertion gain measures. With participants, the audibility of high frequencies, adjustments to prescriptive recommendations in the 4–10 kHz range, and sound quality paired comparisons were completed.

Calculating Real Ear Aided Responses from Insertion Gain and Stimuli for Speech Intelligibility Index and Loudness Modeling

The input ILTASS speech spectrum stimulus was added to the insertion gain values along with the free-field to eardrum transfer function of Moore et al (1997) and ANSI S3.4 (2007) to yield a real-ear aided response (REAR) value. REARs along with hearing threshold levels (HTLs) for each subject were required as input variables for the ANSI S3.5 (1997) SII model as well as the Moore and Glasberg (2004) loudness model.

Audibility of High Frequencies with NAL-NL2 and CAM2

The audibility of a 5, 6.3, and 8 kHz narrowband noise with frequency bandwidth equal to one-third octave was measured using an adaptive threshold procedure approach with 2-down, 1-up (−10 dB +5 dB) step size
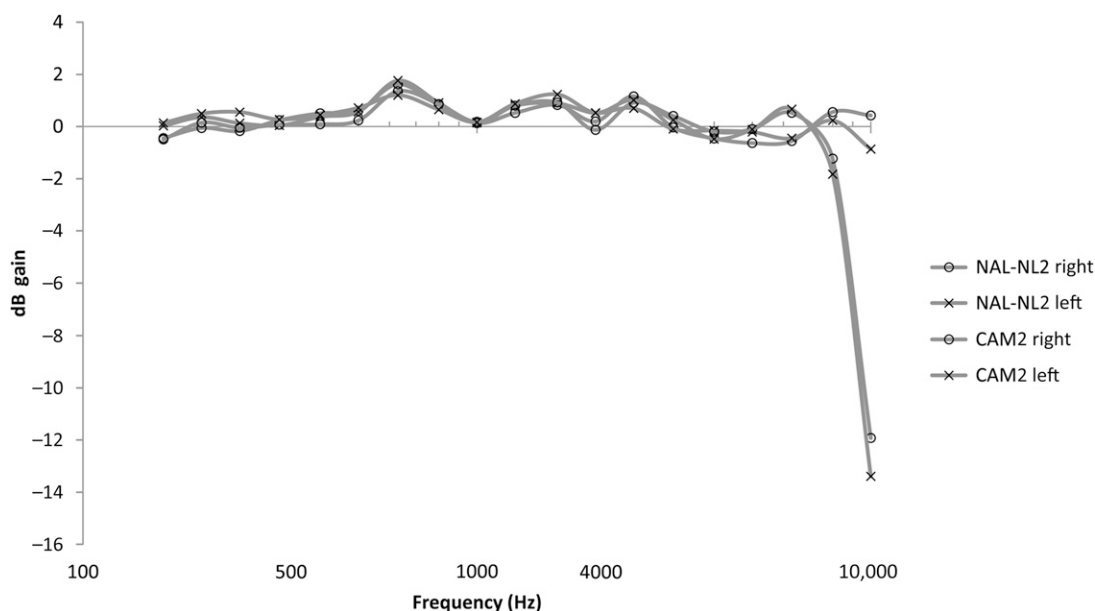
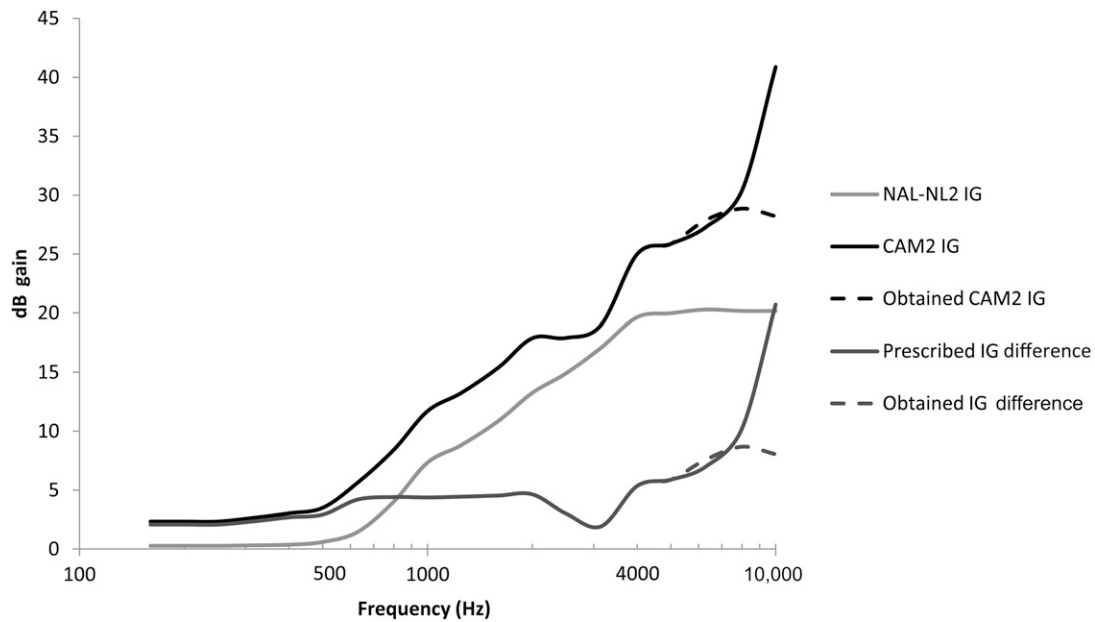**Figure 2.** The obtained match to REIG target of NAL-NL2 and CAM2.

Figure 3. Average frequency-specific differences in the initial prescribed and obtained insertion gain (IG) of NAL-NL2 and CAM2. The IG targets of both prescriptions were met except where noted by the dashed line in the frequency range of 8–10 kHz; obtained IG was 12 dB below the CAM2 target at 10 kHz as shown in Figure 2.

intervals until the first reversal followed by step size adjustments to −2 dB +1 dB. Once the hearing threshold of participants for the narrowband noises was established, these threshold measurements were compared to the prescribed levels of the narrowband noise in the ears of participants. The difference between threshold and prescribed level measurements was termed audibility of the high-frequency stimuli. Audibility was averaged across the left and right ears for simplicity of reporting and because symmetry of hearing loss yielded similar audibility in the two ears of each participant.

Participant Adjustments to Prescriptive Recommendations in the 4–10 kHz Range

Similar in principle to the design of a treble response adjustment, the procedure allowed for participants to adaptively adjust the gain and output in the high frequencies between 4 and 10 kHz. The adjustment process was referenced to the starting prescription and then moved above the prescription by 10 dB and below the prescription by 10 dB followed by a return to the starting prescription. The participants then adjusted the prescription in ±2 dB steps via a thumb up or down cue to the experimenter. If desired, participants could adjust gain further than ±10 dB demonstration range. The adjustment process by each participant was completed three times for subsequent reliability analysis.

In support of this methodology choice, the movement range of ±10 dB when demonstrating changes in gain to participants was selected based on Jenstad et al (2007), who demonstrated that a range of 10 dB in both low

(315–1000 Hz) and high frequencies (1250–4000 Hz) is sufficient to achieve both subjective and objective optimization on a range of outcome measures. The 10 dB range has also been a traditional option for hearing aid volume controls.

Experimental Procedure of Sound Quality Paired Comparison Evaluations

Using a paired comparison technique of the experimental conditions NAL-NL2 versus CAM2, each participant made sound quality judgments for both clarity and pleasantness with experimental stimuli chosen to represent three sound source types: female speech in quiet, female speech in noise, and classical music. At a latter research session, average interval of 23.7 days, participants also made an overall judgment of sound quality preference between the gain-frequency response following adjustment high-frequency gain (4–10 kHz) by each participant and the initial prescription recommendation. Hence, the comparisons were CAM2 versus CAM2 adjusted and NAL-NL2 versus NAL-NL2 adjusted. These comparisons were completed to examine whether the participant adjustment in the 4–10 kHz range from the prescriptive recommendation could be perceptually differentiated, and perhaps preferred, when compared to the initial prescriptive recommendation at a subsequent occasion.

A strength of rating scale was added to accentuate the difference in the sound quality judgments (Keidser et al, 1995; Ricketts and Hornsby, 2005; Johnson et al, 2007; Moore et al, 2011). Participants specified strength

of preferences between paired conditions by expressing which conditions were "much better," "moderately better," and "slightly better" corresponding with assigned ordinal numbers of 3, 2, and 1, respectively, while the nonpreferred memory was assigned a -3, -2, and -1. These number values assigned to degrees of preference should not be confused with an equal distance representation (i.e., interval scale) as perceptual distance for the strength of preference ratings may not be equivalent (Speaks, 1992; Howell, 2002; Johnson et al, 2007). Therefore, nonparametric statistics served as the data analysis method of choice. The comparisons with the strength of preference rating were analyzed with Wilcoxon signed rank test.

RESULTS

Accuracy of the FreqShaping Software Program Gain Adjustments to Real-Ear Gain Changes

Via analysis of the hand-written recorded changes made to the software gain bands during the participant adjustment of the 4–10 kHz bandwidth with the change observed in KEMAR recorded output SPL values, a near perfect 1:1 ratio was evident for all 14 participants. This fact was demonstrated in all one-third octave bands inclusive of 4–10 kHz with a paired samples t-test revealing no significant difference between software gain band changes and the respective changes in KEMAR recorded output SPL values. Regression analysis, using Statistical Package for the Social Sciences (SPSS) 14.0 (2005), of the two variables on a bivariate plot returned an overall correlation value of 0.99 and a

respective 1.0019:1 coefficient multiplier indicating a near perfect 1:1 ratio of software gain band adjustments to real-ear change in SPL.

Reliability of the Participants for High-Frequency Gain Adjustment

The high-frequency gain adjustment process thrice repeated by participants was determined reliable by examining the mean difference between first, second, and third trial adjustments using a repeated measures analysis of variance. No difference was present between trials for NAL-NL2, $F(2,26) = 1.903$, $p = 0.774$, or for CAM2, $F(2,26) = 2.335$, $p = 0.117$. The averaged adjustment was used consequently for subsequent data reporting.

Speech Intelligibility Index and Loudness Modeling

The SII model indicated no clinically meaningful difference in predicted speech recognition between NAL-NL2 and CAM2. This was also true following the adjustments to the prescribed gain in the 4–10 kHz range (Fig. 4). Utilizing a transfer function to percent correct performance from the SII value appropriate for sentence level material (i.e., Humes, 2002), better than 99.5% correct recognition in quiet can be expected for these two prescriptions with patients of similar hearing loss magnitude and configuration.

The loudness model indicated a salient difference between the prescriptions with statistical significance using a paired t-test, $t(13) = 16.579$, $p < 0.001$ (Fig. 5). The NAL-NL2 prescription prescribed an average
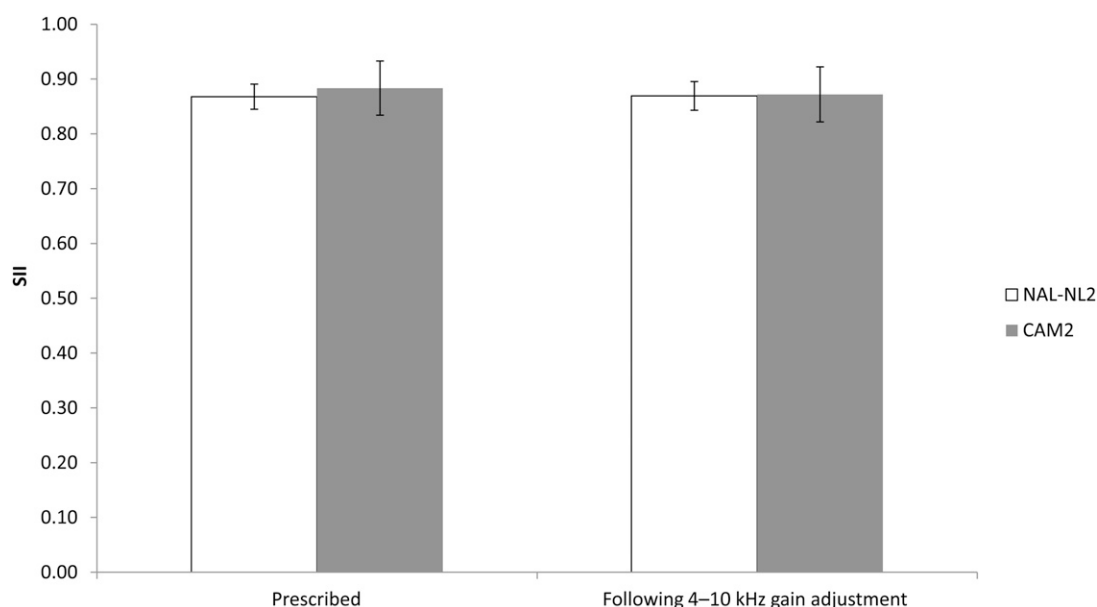

Figure 4. Speech intelligibility index (SII) values of the two prescriptions as prescribed and adjusted were statistically the same and yielded better than 99.5% correct performance for speech in quiet using a transfer function for sentence material in quiet. Error bars are 95% confidence intervals.

144

loudness of 12.2 sones, and CAM2 prescribed an average loudness of 18.3 sones. Expressed to the reference of 12.2 sones from the NAL-NL2 prescription for hearing losses of these participants, CAM2 prescription provided a loudness that was 50% louder. Following adjustment to the 4–10 kHz frequency range by each participant, the respective average loudness was 13.5 and 17.6 sones. This was not a statistically significant change for NAL-NL2, $t(13) = -1.003$, $p = 0.334$, or CAM2, $t(13) = .495$, $p = 0.629$, from the originally prescribed loudness. The nonsignificant effect of the treble adjustment on loudness is mostly reflective of the lack of adjustment control given to participants for changing gain below 4 kHz where frequencies tend to dominate loudness perception.

Prescribed Audibility of High Frequencies with NAL-NL2 and CAM2

There was a recurring statistical difference in the measured audibility of 5, 6.3, and 8 kHz between the prescriptions with paired t-tests, $t(13) = -6.248$, $p < 0.001$, $t(13) = -10.765$, $p < 0.001$, $t(13) = -9.342$, $p < 0.001$. In all cases, NAL-NL2 prescribed less audibility than did CAM2 (Fig. 6).

Experimental Procedure of Sound Quality Paired Comparison Evaluations

The participants were split in their preference, consistent with the SII data and analyses, based on clarity for either NAL-NL2 or CAM2 if a preference was formed at all. That is, when a particular prescription was preferred

for least two of the three listening stimuli, defined as $p < 0.05$ based on Wilcoxon Signed Rank Tests analysis, a significant preference was determined to exist. Four participants had a preference for NAL-NL2, four had a preference for CAM2, and six had a preference for neither.

When participants were asked to rate the prescriptions for pleasantness, the sound quality having the strongest association with overall sound quality evaluations (Preminger and Van Tasell, 1995a, 1995b), eight of the participants had a preference for NAL-NL2, two had a preference for CAM2, and four had no preference. In order to determine if these preferences could be attributed to prescribed loudness and/or high-frequency amplification, an analysis between those participants preferring NAL-NL2 and CAM2 is reported.

Attributing Preferences to Loudness and/or High-Frequency Amplification

The eight subjects preferring NAL-NL2 were prescribed 12.1 sones, and the two subjects preferring CAM2 were prescribed a loudness of 18.8 sones; this difference was statistically significant even with such small group sizes using an independent samples t-test, $t(8) = -4.4$, $p = 0.002$ (similar to Fig. 4). When examining the adjustment to high-frequency gain (4–10 kHz), those participants who preferred NAL-NL2 adjusted the NAL-NL2 gain, on average, by only -1.0 dB; those participants who preferred CAM2 adjusted NAL-NL2 gain, on average, by $+14.3$ dB. This difference was also statistically significant $t(8) = -6.330$, $p < 0.001$ (Fig. 7). Regarding adjustments of CAM2 recommended 4–10 kHz response,
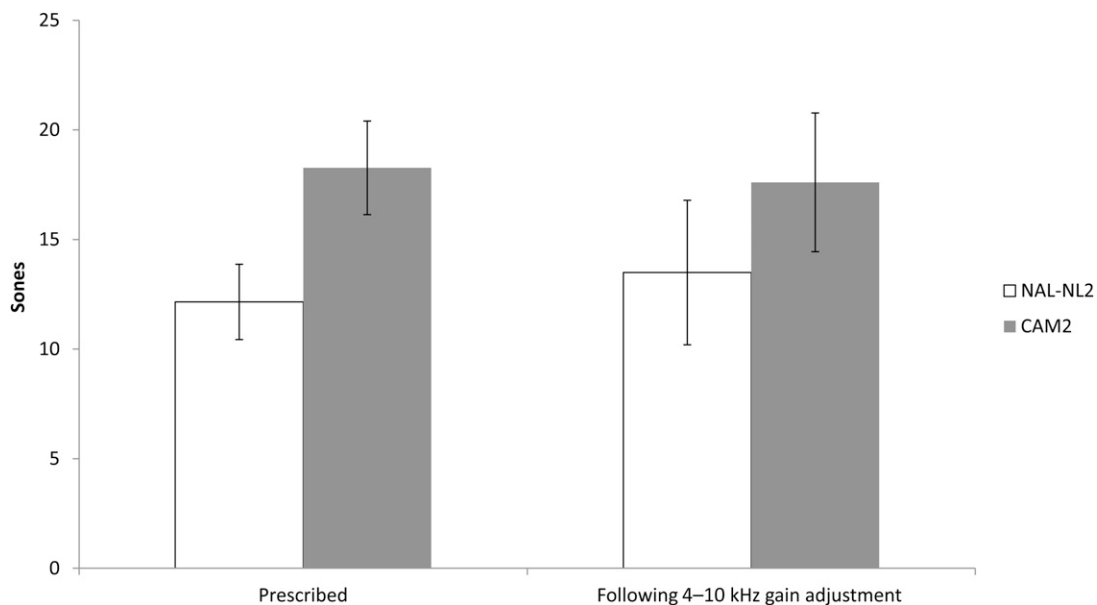

Figure 5. CAM2 is markedly (50%) louder than NAL-NL2; adjustments to the 4–10 kHz frequency region did not statistically alter in a significant manner the average loudness of the fitting. Error bars are 95% confidence intervals.
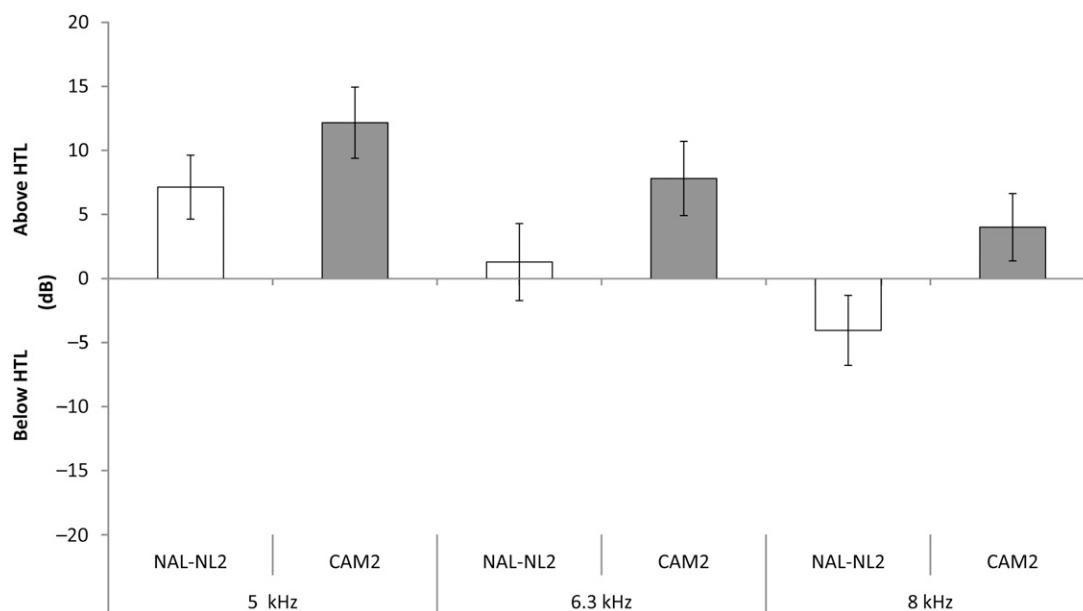
145

Figure 6. Audibility of high-frequency narrowband noises with NAL-NL2 and CAM2; CAM2 prescribed significantly greater audibility than did NAL-NL2. Error bars are 95% confidence intervals.

those participants who preferred NAL-NL2 adjusted CAM2 downward by −6.9 dB; participants preferring CAM2 adjusted CAM2 gain, on average, by only −1.1 dB. This difference too was statistically significant at t(8) = −3.581, p = 0.009 with unassumed equal variances (Fig. 7).

The adjustments to the recommended 4–10 kHz responses from both NAL-NL2 and CAM2 across all 14 subjects were also analyzed. That is, prescribing gain in the 4-10 kHz range is novel to prescriptive approaches and foundational research for a recommendation is limited at best. When averaging across subjects the 4-10 kHz gain of NAL-NL2 was increased by 1.8 dB and CAM2 was decreased by 3.9 dB, a statistically significant difference t(13) = -5.535, p = 0.001. To determine whether these adjustments were different from the initial prescriptive recommendation, a one sample t-test from a test value of 0 dB was completed. The +1.8 dB average adjustment to NAL-NL2 by the 14 participants was not significantly different from 0 dB,
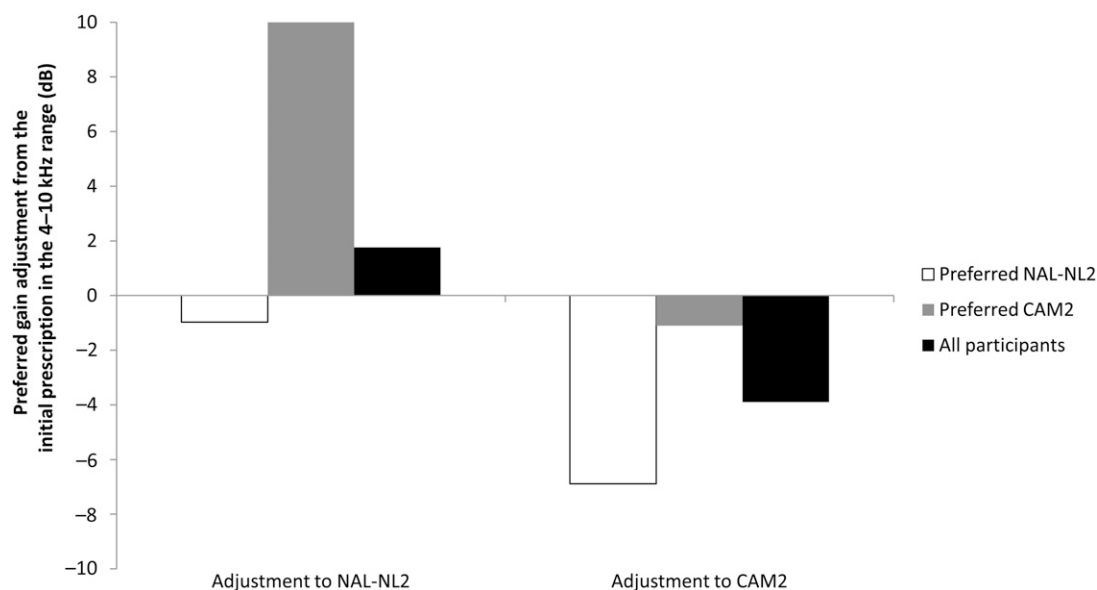

Figure 7. Adjustments to the NAL-NL2 and CAM2 prescriptions in the frequency range of 4–10 kHz by study participants. The following were statistically significant differences: (1) participants preferring CAM2 adjusted NAL-NL2 gain upwards more so than participants preferring NAL-NL2, (2) participants preferring NAL-NL2 adjusted CAM2 gain downward more so than participants preferring CAM2, (3) across all participants CAM2 gain was adjusted downward while NAL-NL2 gain was not; the turned down amount of ~4 dB for CAM2 was significantly different from zero (or the initial prescription).

146

t(13) = 1.113, p = 0.286. The -3.9 dB average adjustment to CAM2 by the 14 participants was significantly different from 0 dB, t(13) = -2.285, p = 0.04. These results indicate the participants did not make adjustments that significantly deviated from the initial recommendation of NAL-NL2 but did from the recommendations of CAM2 in the range of 4-10 kHz (Fig. 7).

New versus Experienced Hearing Aid Users

Due to the inclusion of new and experienced hearing aid users, an analysis of the data to determine whether such experience impacted the adjustment of high-frequency gain was completed. A two-way univariate analysis of variance of hearing aid experience and prescription indicated that the adjustment to gain in the 4–10 kHz range could not be differentiated based on new or experienced hearing aid user status, F(1,24) = 0.070, p = .794. Consistent with previous results (Fig. 7), the average downward high-frequency adjustment made to CAM2 were significantly different from the limited upward adjustment to NAL-NL2 targets, F(1,24) = 4.431, p = 0.046. No interaction effect was present, F(1,27) = 1.779, p = .195. These significant differences, p = 0.036, in gain adjustments for NAL-NL2 and CAM2 were the same with hearing loss attributes as covariates including the .5, 1, and 2 kHz PTA, the 2, 4, and 8 kHz PTA, as well as the 3–12 kHz dB per octave slope.

Adjusted High-Frequency Amplification and Preference at a Later Evaluation Session

Lastly, whether the high-frequency gain adjustments to the prescriptions were indeed preferred relative to the initial prescription, at a subsequent session (average 23.7 days from the first session) was determined. This determination included an analysis of the paired comparison of overall sound quality including both the perceptual attributes of clarity and pleasantness according to the instruction card provided to the participant. The overall sound quality comparisons of *NAL-NL2* versus *NAL-NL2 adjusted* and *CAM2* versus *CAM2 adjusted* comparisons were analyzed with determination of a statistical preference applied as described previously.

Gain deviations between the occurrences when a preference was established and occurrences when a preference was not established yielded a significance difference, t(26) = 2.183, p = 0.038 (Fig. 8). The data suggest that when a preference for adjusted settings existed, a difference of greater than 4 dB between the initial prescription and adjusted prescription was likely to be expected. Analyzing the 13 occurrences in which the adjusted settings were more than 4 dB from the initial prescription, nine of those times (69%) the participants preferred the adjusted settings over the initial prescription.
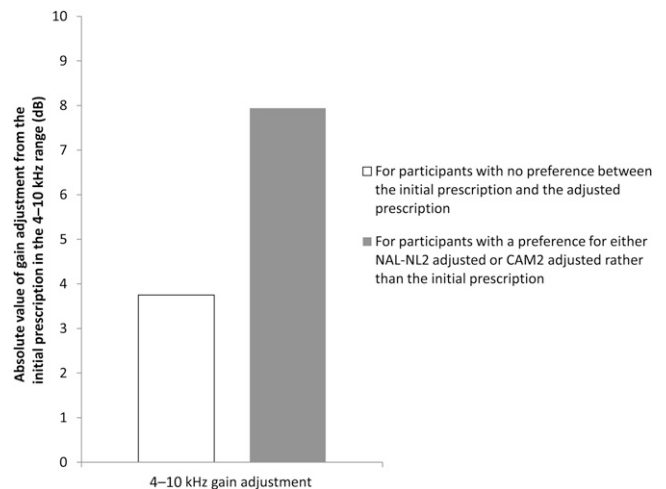

Figure 8. For those instances in which participants had a preference for adjusted gain settings over the initial prescription (gray bar), there was significantly more adjustment to gain (greater than 4 dB) by the participants than when participants had no preference for the adjusted prescription over the initial prescription (white bar).

DISCUSSION

These results show clear differences between the initial prescriptive recommendations of hearing aid amplification for adults by the NAL-NL2 and CAM2 methods. Insertion gain differences create substantial loudness differences yet comparable predicted speech intelligibility. Also evident was the fact that CAM2 did prescribe more audibility to higher frequencies than did NAL-NL2. Such evidence is a reflection of each prescription's implementation.

As hearing loss becomes more severe, NAL-NL2 intentionally chooses to prescribe less audibility or no audibility based on assumed desensitization (reduced utility) for the individual (inferred from large-scale group data of individuals with similar hearing loss levels); whereas, CAM2 attempts to restore audibility at 10 kHz by design no matter the severity of hearing loss. The CAM2 recommendations are limited, however, by the ability of hearing aids to actually achieve its gain targets which have proved difficult beyond 8 kHz not only in this study but also with commercial instruments in the study by Füllgrabe et al (2010). CAM2 targets do confirm nonetheless that restoring the RMS of speech in the high-frequencies to audible levels requires considerable gain. In cases were no audibility can be achieved for hearing aids fit to either NAL-NL2 or CAM2 amplitude compression targets, the need for use of technologies such as frequency lowering strategies is apparent and, by inferred logic, perhaps necessary.

Data collected in this study further support the finding that most individuals with sensorineural hearing loss will prefer a hearing aid prescription that yields a loudness much below that of the listener with normal hearing sensitivity. This agrees with previous generation prescriptions, like NAL-NL1, which were designed with a

loudness constraint of 23.43 sones (the loudness of a 65 dBA ILTASS by the listener with normal hearing sensitivity using the Moore and Glasberg (1997) loudness model) (Byrne et al, 2001). A conglomeration of studies indicated, however, that NAL-NL1 was still too loud for 46% of listeners, just right for 49% of listeners, and too soft for only 5% (Keidser and Dillon, 2006). The results suggested then that by a reduction in loudness more listeners would be pleased with the gain recommendation NAL-NL1.

As a result, developers of NAL-NL2 set out to reduce its prescribed loudness but with simultaneous use of a revised SII model hopefully achieve comparable speech intelligibility to that of NAL-NL1. A design goal of NAL-NL2 was also to further individualize the recommended amplification based on characteristics such as age and gender of the patient as well as whether the patient had previous hearing aid experience or primarily heard a tonal or nontonal language. The results of Johnson and Dillon (2011) confirmed the reduction in loudness with comparable speech intelligibility and Keidser et al (2011) discusses the adjustments to gain based characteristics of the patient. One reasonable question is how can a reduction in loudness be achieved with comparable speech intelligibility. An explanation is based on the principles of asset allocation – maximizing return and minimizing tradeoffs (risks). The NAL-NL2 is allocating the asset of prescribed gain to maximize speech intelligibility and minimize contributions to loudness.

The CAM2 method provides loudness that is less than the 23.43 sones. However, this finding needs to be interpreted based on the revision of loudness models which asserts the perceived loudness of a normal hearer. That is, the 23.43 sones amount is based on the Moore and Glasberg (1997) model whereas the Moore and Glasberg (2004) model asserts a loudness of the ILTASS at the eardrum given a 65 dB SPL diffuse sound field input was 18.6 sones, agreeing well with calculated loudness of the revised speech spectrum of Moore et al (2008), whose loudness of a 65 dB SPL input is 20.7 sones (Moore et al, 2010). The average estimate of perceived loudness for a normal hearer is hence about 19.65 sones depending upon the assumed speech spectrum. For the audiograms of participants in this study, CAM2 prescribed an average loudness of 18.3 sones, a value that is quite similar to the loudness of a normal hearer. For other audiograms, the loudness of the CAM2 prescription has also been reported as similar to that of NAL-NL1 (Johnson and Dillon, 2011).

Because of the milder hearing losses in this study compared to Johnson and Dillon (2011), the mean loudness of the NAL-NL2 method was higher (12.2 sones) than the mean 8.5 sones reported there. Likewise, the discussion of that manuscript indicated that for severe-to-profound hearing loss, the differences between NAL-NL2 and CAM2 will likely be larger because of NAL's use of a revised desensitization factor that is more pronounced than the one described in ANSI S3.5 (1997). This revised

factor encourages the use of less gain as hearing loss increases because of the expected reduction in effectiveness of audibility to render comparable improvements to speech intelligibility. In opposition, for milder hearing losses, the factor would allow more gain, and subsequent loudness, predicting useful contributions to speech intelligibility by increased audibility.

The lack of clear preferences or a split preferences among participants between NAL-NL2 and CAM2 on the sound quality attribute of clarity was consistent with the SII data and analyses. That is, clarity is generally correlated best with how clear the speech is to understand and measures of intelligibility (e.g., Gabrielsson and Sjogren, 1979a, 1979b; Gabrielsson et al, 1988; Eisenberg et al, 1997; Eisenberg et al, 1998). Because of the softer loudness prescribed by NAL-NL2 which presumably lead to a more pleasantness, the majority of participants in this study expressed an initial preference for the prescription when compared to CAM2. For those patients with preference for a louder fitting, CAM2 was a their choice.

Some patients will prefer slight deviations in gain from prescribed settings at the initial fitting which are still preferred at a later date; in this study when the adjustment from the prescription between 4–10 kHz exceeded 4 dB from either NAL-NL2 or CAM2 the participants demonstrated a repeated preference for that adjustment 69% of the time; hence, not all patients demonstrated consistent preferences for adjusted frequency responses consistent with Keidser et al (2008). These findings are viewed as limited evidence for allowing adult patients, capable of doing so, to adjust the initially prescribed amplification of their hearing via training implementations and interfaces.

To clarify the term *training,* it used here and elsewhere in commercial hearing aid amplification as the ability to adjust gain for multiple input levels (compression training) as well as at certain frequencies (frequency response training). A few examples in modern commercial hearing aids of training are Self Learning by Unitron, User Preference Tuning by Phonak, SoundLearning by Siemens, and Self Learning by Starkey. All of which are compression parameter adjustments via analysis of input level and patient changes to the volume control. An example of a treble frequency response adjustment to the hearing aid frequency response is the remote control feature of SoundBalance by Siemens.

The finding that hearing aid user status of new versus experienced did not impact the high-frequency gain adjustments in this study may be for the following reason. Both the NAL-NL2 and CAM2 prescriptions provided more high-frequency gain than the NAL-NL1 prescription of a 6 kHz bandwidth with which the experienced participants in this study would have had listening familiarity. Hence when adjusting gain in the 4-10 kHz bandwidth, both the new and experienced hearing aid participants would have been receiving amplification of a frequency range that was newly

audible again or, at least, had greater audibility than the participants were used to hearing.

To reiterate, these study results are based on participant judgments with linear gain processing from prescribed gains for an average level input with nonlinear gain processing prescriptions. Given the general purpose of nonlinear processing, however, is to offer greater amplification for improved audibility of soft input levels and less amplification for improved comfort of high input levels, both types of gain processing offer comparable gain for an average input level (e.g., Byrne, 1996; Byrne, et al, 2001; Dillon, 2001). Also because linear gain processing was utilized, the effect on sound quality preferences for variable attack and release compression time constants that exist across manufacturers of hearing aids in the marketplace was not evaluated in this study. This may be noteworthy concern when considering these study results for possible extrapolation to the clinical environment. Such an examination would add considerable complexity and scope to the current study; more importantly, it should be acknowledged that generic prescriptive methods of both NAL-NL2 and CAM2 do not make recommendations of gain based on compression time constants for manufacturer-specific compression architectures. Nor does either prescription recommend specific compression time constants. Hence, there is not empirical evidence to expect that linear gain processing is not sufficient for evaluating gain differences between the NAL-NL2 and CAM2 prescriptions for an average input level of speech and music stimuli. It is true that this study can assert no ability to report on sound quality differences between the two prescriptions for soft and louder input levels and has not done so.

In summary, allowing patients to compare validated hearing aid prescriptions at the time of the initial fitting is another step toward enabling the patient to direct his/her treatment in within a range of reasonable options. The comparisons could be accomplished by fitting each prescription in a program (memory) of the hearing aid and proceeding in a manner similar to that discussed herein. An area of further research to find optimal test stimuli and a paradigm for comparisons is currently underway in an open-fit hearing aid configuration incorporating nonlinear gain processing. The implementation of such comparisons would bring hearing aid fitting processes full circle back to the comparative approach (Carhart, 1946). The premise of such comparisons, however, are not expected to be based on finding the hearing aid prescription providing the best speech recognition necessarily but on identifying the prescription that is initially preferred by the patient which has already been scientifically validated to provide good patient outcomes in a number of areas. The premise of comparison is intuitively imperative in the recently began era of allowing patients, able to do so, to train the amplification characteristics of their hearing aids from essentially day one. Rationale for finding the "best" starting point for training is supported by research indicating that when listeners are allowed to train the frequency- and input-dependent gain of the hearing aids, the end result is affected by the starting response (baseline) (Dreschler et al, 2008; Keidser et al, 2008; Mueller et al, 2008).

REFERENCES

American National Standards Institute (ANSI). (1997) *Methods for Calculation of the Speech Intelligibility Index*. ANSI S3.5-1997. New York: American National Standards Institute.

Byrne D. (1996) Hearing aid selection for the 1990s: where to? *J Am Acad Audiol* 7:377–395.

Byrne D, Dillon H, Ching T, Katsch R, Keidser G. (2001) NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *J Am Acad Audiol* 12:37–51.

Byrne D, Dillon H, Tran K, et al. (1994) An international comparison of long-term average speech spectra. *J Acoust Soc Am* 96: 2108–2120.

Carhart R. (1946) Tests for the selection of hearing aids. *Laryngoscope* 56:780–794.

Cornelisse L, Seewald RC, Jamieson DG. (1995) The input/output formula: a theoretical approach to the fitting of personal amplification. *J Acoust Soc Am* 97:1854–1864.

Cox RM. (1995) Using loudness data for hearing aid selection: the IHAFF approach. *Hear J* 48:10–44.

Cox RM, Alexander G, Gilmore C. (1987) Development of the Connected Speech Test (CST). *Ear Hear* 8:119S–126S.

Dillon H. (2001) *Hearing Aids*. Sydney, Australia: Boomerang Press.

Dillon H, Ching T, Keidser G, Katsch R. (2008) Derivation of the NAL-NL2 prescription formula. Paper presented at Audiology NOW!, Charlotte, NC.

Dirks D, Kincaid G. (1987) Basic acoustic considerations of ear canal probe measurements. *Ear Hear* 8:60S–67S.

Dreschler W, Keidser G, Convery E, Dillon H. (2008) Client-based adjustments of hearing aid gain: the effect of different control configurations. *Ear Hear* 29:214–227.

Eisenberg L, Dirks D, Gornbein J. (1997) Subjective judgments of speech clarity measured by paired comparisons and category rating. *Ear Hear* 18:294–306.

Eisenberg L, Dirks D, Takayangi S. (1998) Subjective judgments of clarity and intelligibility for filtered stimuli with equivalent speech intelligibility index predictions. *J Speech Lang Hear Res* 41:327–339.

Fletcher H, Munson WA. (1933) Loudness, its definition, measurement and calculation. *J Acoust Soc Am* 5:82–108.

Füllgrabe C, Baer T, Stone M, Moore B. (2010) Preliminary evaluation of a method for fitting hearing aids with extended bandwidth. *Int J Audiol* 49:741–753.

Gabrielsson A, Schenkman BN, Hagerman B. (1988) The effects of different frequency responses on sound quality judgments and speech intelligibility. *J Speech Hear Res* 31:166–177.

Gabrielsson A, Sjogren H. (1979a) Perceived sound quality of hearing aids. *Scand Audiol* 8:159–169.

Gabrielsson A, Sjogren H. (1979b) Perceived sound quality of sound-reproducing systems. *J Acoust Soc Am* 65:1019–1033.

Harris RW, Goldstein DP. (1985) Hearing aid quality judgments in reverberant environments using a magnitude estimation procedure. *Audiology* 24:32–43.

Hornsby BWY, Johnson EE, Picou E. (2011) Effects of degree and configuration of hearing loss on the contribution of high- and low-frequency speech information to bilateral speech understanding. *Ear Hear* 32:543–555.

Howell DC. (2002) *Statistical Methods for Psychology*. 5th edition. Belmont, CA: Duxbury Press.

Humes L. (2002) Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *J Acoust Soc Am* 112:1112–1132.

Humes LE. (2003) Modeling and predicting hearing aid outcome. *Trends Amplif* 7(2):41–75.

Humes L, Hackett T. (1986) An evaluation of several rationales for selecting hearing aid gain. *J Speech Hear Disord* 51:272–281.

Jenstad LM, Bagatto MP, Seewald RC, Scollie SD, Cornelisse LE, Scicluna R. (2007) Evaluation of the desired sensation level [input/output] algorithm for adults with hearing loss: the acceptable range for amplified conversational speech. *Ear Hear* 28:793–811.

Johnson EE, Dillon H. (2011) A comparison of gain for adults from generic hearing aid prescriptive methods: impacts on predicted loudness, frequency bandwidth, and speech intelligibility. *J Am Acad Audiol* 22:441–459.

Johnson EE, Ricketts T, Hornsby B. (2007) Effects of digital feedback reduction algorithms in modern digital hearing aids on sound quality. *J Am Acad Audiol* 18:410–416.

Keidser G, Dillon H. (2006) What's new in prescriptive fittings Down Under. In: Palmer C, Seewald R, eds. *Hearing Care for Adults*. Stafa, Switzerland: Phonak AG, 133–142.

Keidser G, Dillon H, Byrne D. (1995) Candidates for multiple frequency response characteristics. *Ear Hear* 16:562–574.

Keidser G, Dillon H, Convery E. (2008) The effect of the base line response on self-adjustments of hearing aid gain. *J Acoust Soc Am* 124:1668–1681.

Keidser G, Dillon H, Flax M, Ching T, Brewer S. (2011) The NAL-NL2 prescription procedure. *Audiol Res* 1:88–90.

Killion M. (1994) *Fig6.exe Software: Hearing Aid Fitting Targets for 40, 65 & 95 dB SPL Inputs. (Version 1.01D)*. Elk Grove Village, IL: Etymotic Research.

Killion M. (1995) Loudness-data basis for "FIG6" hearing-aid fitting targets. *J Acoust Soc Am* 98:2927.

Kochkin S. (2011) MarkeTrak VIII: reducing patient visits through verification and validation. *Hear Rev* 18:10–12.

Logan SA, Schwartz DM, Ahlstrom JB, Ahlstrom C. (1984) Acoustical environmental influences on hearing-aid quality/intelligibility judgments. *ASHA* 26:110.

Moore BCJ, Füllgrabe C. (2010) Evaluation of the CAMEQ2-HF method for fitting hearing aids with multichannel amplitude compression. *Ear Hear* 31:657–666.

Moore BCJ, Füllgrabe C, Stone MA. (2011) Determination of preferred parameters for multichannel compression using individually fitted simulated hearing aids and paired comparisons. *Ear Hear* 32:556–568.

Moore BCJ, Glasberg BR. (1997) A model of loudness perception applied to cochlear hearing loss. *Aud Neurosci* 3:289–311.

Moore BCJ, Glasberg BR. (2004) A revised model of loudness perception applied to cochlear hearing loss. *Hear Res* 188:70–88.

Moore BCJ, Glasberg BR, Baer T. (1997) A model for the prediction of thresholds, loudness and partial loudness. *J Audio Eng Soc* 45:224–240.

Moore BCJ, Glasberg BR, Stone MA. (2010) Development of a new method for deriving initial fitting for hearing aids with multi-channel compression: CAMEQ2-HF. *Int J Audiol* 49:216–227.

Moore BCJ, Stone MA, Füllgrabe C, Glasberg BR, Puria S. (2008) Spectro-temporal characteristics of speech at high frequencies and the potential for restoration of audibility to people with mild-to-moderate hearing loss. *Ear Hear* 29:907–922.

Mueller HG, Hornsby BWY, Weber JE. (2008) Using trainable hearing aids to examine real-world preferred gain. *J Am Acad Audiol* 19:758–773.

Preminger JE, Van Tasell DJ. (1995a) Quantifying the relation between speech quality and speech intelligibility. *J Speech Hear Res* 38:714–725.

Preminger JE, Van Tasell DJ. (1995b) Measurement of speech quality as a tool to optimize the fitting of a hearing aid. *J Speech Hear Res* 38:726–736.

Rankovic CM. (1991) An application of the articulation index to hearing aid fitting. *J Speech Hear Res* 34:391–402.

Ricketts T, Hornsby B. (2005) Sound quality measures for speech in noise through a commercial hearing aid implementing digital noise reduction. *J Am Acad Audiol* 16:272–279.

Ricketts T, Van Vliet D. (1996) Updating our fitting strategies. *Am J Audiol* 5:29–35.

Scollie S, Seewald R, Cornelisse L, et al. (2005) The desired sensation level multistage input/output algorithm. *Trends Amplif* 9:159–197.

Seewald RC, Cornelisse LE, Ramji KV, Sincliar ST, Moodie KS, Jamieson DG. (1996) *DSL v4.0 for Windows Manual*. London, ON: University of Western Ontario.

Smeds K, Keidser G, Zakis J, et al. (2006a) Preferred overall loudness. I. Sound field presentation in the laboratory. *Int J Audiol* 45:2–11.

Smeds K, Keidser G, Zakis J, et al. (2006b) Preferred overall loudness. II. Listening through hearing aids in field and laboratory tests. *Int J Audiol* 45:12–25.

Speaks C. (1992) *Introduction to Sound: Acoustics for the Hearing and Speech Sciences*. San Diego: Singular.

Valente M, Van Vliet D. (1997) The independent hearing aid fitting forum (IHAFF) protocol. *Trends Amplif* 2:6–35.

Valente M, Abrams H, Benson D, Chisolm T, Citron D, Hampton D, Loavenbruck A, Ricketts T, Solodar H, Sweetow R. (2006) Guidelines for the audiologic management of adult hearing impairment. *Audiol Today* 18:32–37.